

CLAIMS

1. (Previously Presented) A method, comprising:

receiving an identifier of a first participant of a plurality of participants of a multi-party conference;

receiving a first signal indicating that the first participant is providing information to the multi-party conference; and

displaying, on a user interface, a first indication that the first participant is providing information to the multi-party conference in response to the received first signal, wherein the first indication comprises displaying, on the user interface, the identifier at a top position within a list of identifiers of the plurality of participants, the position designated for a participant providing information to the multi-party conference wherein the identifier of the first participant comprises a location associated with the first participant.

2. (Original) The method of claim 1, further comprising terminating the first signal when the first participant is no longer providing information to the multi-party conference.

3. (Original) The method of claim 2, further comprising terminating display of the first indication in response to termination of the first signal.

4. (Original) The method of claim 1, further comprising:
while displaying the first indication, displaying an identifier of a second participant of the plurality of participants without indicating that the second participant is providing information to the multi-party conference.
5. (Original) The method of claim 1, wherein the first indication comprises displaying the identifier so as to flash.
6. (Original) The method of claim 1, wherein the first indication comprises displaying the identifier using a preselected color scheme.
7. (Original) The method of claim 1, wherein the first indication comprises displaying the identifier as an animated graphic.
8. (Original) The method of claim 1, wherein the first indication comprises displaying the identifier using video.
9. (Original) The method of claim 1, wherein the first indication comprises displaying the identifier as a photograph.
10. (Original) The method of claim 1, wherein the first indication comprises displaying the identifier with a graphic proximate to the identifier of the first participant.

11. (Original) The method of claim 10, wherein the graphic proximate to the identifier of the first participant comprises text.

12. (Previously Canceled)

13. (Previously Canceled)

14. (Previously Canceled)

15. (Original) The method of claim 1, wherein the first signal is received when the first participant is speaking over an audio link.

16. (Original) The method of claim 15, wherein the audio link comprises a telephone link.

17. (Original) The method of claim 15, wherein the audio link comprises a satellite link.

18. (Original) The method of claim 15, wherein the audio link comprises a voice over Internet Protocol (VoIP) link.

19. (Original) The method of claim 1, further comprising receiving information from a second participant of the plurality of participants via a network while the first participant is providing information to the multi-party conference.

20. (Original) The method of claim 19, wherein receiving information from the second participant further comprises receiving information via client-server communication over the network.

21. (Original) The method of claim 19, wherein receiving information from the second participant further comprises receiving information via peer-to-peer communication over the network.

22. (Original) The method of claim 19, wherein the network comprises the Internet.

23. (Original) The method of claim 19, wherein the network comprises a local area network.

24. (Original) The method of claim 19, wherein the network comprises a wireless network.

25. (Original) The method of claim 1, further comprising:

concurrently with the first indication, displaying a second indication that a second participant of the plurality of participants is providing information to the multi-party conference in response to receiving a second signal concurrently with the first signal, the second signal indicating that the second participant is providing information to the multi-party conference.

26. (Previously Presented) The method of claim 1, further comprising suppressing an indication of second participant of the plurality of participants, in response to receiving a signal concurrently indicating that the second participant is providing information to the multi-party conference.

27. (Previously Presented) A computer-readable tangible medium having computer-executable instructions that, upon execution, facilitate a computing device in performing operations comprising:

receiving an identifier of a first participant of a plurality of participants of a multi-party conference;

receiving a first signal indicating that the first participant is providing information to the multi-party conference; and

displaying, by a user interface configured to display, a first indication that the first participant is providing information to the multi-party conference in response to the received first signal, wherein the first indication comprises displaying, on the user interface, the identifier at a top position within a list of identifiers of the plurality of participants, the position designated for a participant providing information to the multi-

party conference wherein the identifier of the first participant comprises a location associated with the first participant.

28. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the operations further comprise:

while displaying the first indication, displaying an identifier of a second participant of the plurality of participants without indicating that the second participant is providing information to the multi-party conference.

29. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the first indication comprises displaying the identifier so as to flash.

30. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the first indication comprises displaying the identifier using a preselected color scheme.

31. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the first indication comprises displaying the identifier as an animated graphic.

32. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the first indication comprises displaying the identifier using video.

33. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the first indication comprises displaying the identifier as a photograph.

34. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the first indication comprises displaying the identifier with a graphic proximate to the identifier of the first participant.

35. (Previously Presented) The computer-readable tangible medium of claim 34, wherein the graphic proximate to the identifier of the first participant comprises text.

36. (Previously Canceled)

37. (Previously Canceled)

38. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the first signal is received when the first participant is speaking over an audio link.

39. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the operations further comprise:

concurrently with the first indication, displaying a second indication that a second participant of the plurality of participants is providing information to the multi-party conference in response to receiving a second signal concurrently with the first signal,

the second signal indicating that the second participant is providing information to the multi-party conference.

40. (Previously Presented) The computer-readable tangible medium of claim 27, wherein the operations further comprise:

suppressing an indication of second participant of the plurality of participants, in response to receiving a signal concurrently indicating that the second participant is providing information to the multi-party conference.

41. (Previously Presented) A system for multi-party conferencing for a plurality of participants, the system comprising:

a user interface (UI); and

an information source indicator to cause the UI to display an identifier corresponding to a participant in response to receiving a signal that indicates that the participant is providing information to the multi-party conference, wherein the displayed identifier is at a position within a list of identifiers of the plurality of participants, the position designated for a participant providing information to the multi-party conference wherein the identifier corresponding to the participant comprises a location associated with the participant.

42. (Original) The system of claim 41, wherein the signal terminates in response to the participant no longer providing information to the multi-party conference.

43. (Original) The system of claim 42, wherein the information source indicator is to cause the UI to display the identifier without the property in response to the signal terminating.

44. (Original) The system of claim 41, wherein the information source indicator is further to cause the UI display a plurality of identifiers corresponding to the plurality of participants without the property that indicates that the plurality of corresponding participants are providing information to the multi-party conference.

45. (Original) The system of claim 41, wherein the information provided by the participant is the participant's speech over an audio link.

46. (Original) The system of claim 41, wherein the UI is to selectively display additional information received from the network while the participant is providing information to the multi-party conference.

47. (Previously Presented) A method, comprising:

via a first participant equipment:

detecting an acoustic signal;

determining whether the detected acoustic signal was generated by a person speaking by receiving a frame of audio data derived from the detected acoustic signal;

classifying the received frame based on spectral data of the received frame, the spectral data obtained by performing a modulated complex lapped transform (MCLT) on the frame of audio data, the classifying comprising classifying the received frame as one of the plurality of predetermined frame types comprising a live-type frame, a phone-type frame, and an unsure-type frame, wherein live-type frames represent frames determined to be derived from acoustic signals generated by a person speaking, and phone-type frames represent frames determined to be derived from acoustic signals generated by an audio transducer device; and

providing a signal indicating to a second participant equipment that the detected acoustic signal was generated by the person.

48. (Original) The method of claim 47, wherein determining whether the detected acoustic signal was generated by a person further comprises:

determining whether the detected acoustic signal was speech from an audio transducer device.

49. (Original) The method of claim 47, wherein the signal is transmitted over a network to participants of a multi-party conference.

50. (Previously Presented) The method of claim 47 wherein determining whether the detected acoustic signal was generated by a person further comprises:

determining a source of a portion of the detected acoustic signal used to derive the frame based on the classification of the frame and a prior determination of a source of a detected acoustic signal portion.

51. (Original) The method of claim 50, wherein the spectral data is obtained by performing a frequency transform on the frame of audio data.

52. (Original) The method of claim 51, wherein the spectral data is obtained by performing a fast Fourier transform (FFT) on the frame of audio data.

53. (Canceled)

54. (Original) The method of claim 50, further comprising determining a first frequency band's energy and a second frequency band's energy from the spectral data.

55. (Original) The method of claim 54, wherein the first frequency band corresponds to a frequency range for consonants and the second frequency band corresponds to a frequency range for vowels.

56. (Original) The method of claim 55, further comprising classifying the frame as being generated by a person when the ratio of the energies of the first and second frequency bands exceeds a first predetermined threshold.

57. (Original) The method of claim 56, further comprising selectively classifying the frame as being from a different source when the ratio of the energies of the first and second frequency bands is below a second predetermined threshold.

58. (Original) The method of claim 57, further comprising selectively classifying the frame as having an unknown source when the ratio of the energies of the first and second frequency bands exceeds the second predetermined threshold and is below the first predetermined threshold, the second predetermined threshold being less than the first predetermined threshold.

59. (Original) The method of claim 55, further comprising determining whether the frame was derived from speech, wherein speech includes acoustic signals generated by a person speaking and acoustic signals generated by an audio transducer device.

60. (Original) The method of claim 59, further comprising determining noise floor energies of the first and second frequency bands using the spectral data, wherein a frame is selectively classified as being derived from speech in response to the energy of the first frequency band exceeding the noise floor energy of the first frequency band or the energy of the second frequency band exceeding the noise floor of the second frequency band, or both.

61. (Original) The method of claim 60, wherein classifying the received frame further comprises:

determining whether a frame received within a predetermined number of frames relative to the received frame has substantially all of its energy in the second frequency band.

62. (Original) The method of claim 60, wherein classifying the received frame further comprises:

determining whether a frame adjacent to the received frame was classified as derived from speech.

63. (Previously Canceled).

64. (Previously Presented) The method of claim 47, further comprising determining the source of the detected acoustic signal to be an acoustic signal generated by a person, if:

the prior determination of a source of a detected acoustic signal portion is that the source was an audio transducer device;

the frame is classified as a live-type frame; and

a predetermined number of prior frames includes live-type frames that exceed a predetermined live-type frame count threshold.

65. (Previously Presented) The method of claim 47, further comprising determining the source of the detected acoustic signal to be unsure, if:

the prior determination of a source of a detected acoustic signal portion is that the source was an audio transducer device;

the frame is classified as a live-type frame;

a predetermined number of most recent frames do not include enough live-type frames to exceed a predetermined threshold; and

an elapsed time since receiving a previous frame derived from speech exceeds a predetermined first time threshold.

66. (Previously Presented) The method of claim 47, further comprising determining the source of the detected acoustic signal to be an audio transducer device, if:

the prior determination of a source of a detected acoustic signal portion is that the source was an acoustic signal generated by a person speaking;

the frame is classified as a phone-type frame;

an elapsed time since receiving a previous live-type frame exceeds a predetermined second time threshold; and

a counter value does not exceed a predetermined count threshold, the counter value to track a number of consecutive non-live-type frames received after receiving a live-type frame of most recent frames do not include enough live-type frames to exceed a predetermined threshold.

67. (Previously Presented) The method of claim 47, further comprising determining the source of the detected acoustic signal to be unsure, if:

the prior determination of a source of a detected acoustic signal portion is that the source was an acoustic signal generated by a person speaking;

the frame is classified as a phone-type frame;

an elapsed time since receiving a previous live-type frame exceeds a predetermined second time threshold; and

the counter value is below a predetermined count threshold, the counter value to track a number of consecutive non-live-type frames received after receiving a live-type frame of most recent frames do not include enough live-type frames to exceed a predetermined threshold.

68. (Previously Presented) The method of claim 47, further comprising determining the source of the detected acoustic signal to be an acoustic signal generated by a person speaking, if:

the prior determination of a source of a detected acoustic signal portion is unsure;

the frame is classified as a live-type frame; and

a predetermined number of most recent prior frames includes live-type frames that exceed in number a predetermined live-type frame count threshold.

69. (Previously Presented) The method of claim 47, further comprising determining the source of the detected acoustic signal to be an acoustic transducer device, if:

the prior determination of a source of a detected acoustic signal portion is unsure;

the frame is classified as a phone-type frame; and

a predetermined number of most recent prior frames includes phone-type frames that exceed in number a predetermined phone-type frame count threshold.

70. (Previously Presented) A computer-readable tangible medium having computer-executable instructions that, upon execution, facilitate a computing device in performing operations comprising:

detecting an acoustic signal;

determining whether the detected acoustic signal was generated by a person speaking by receiving a frame of audio data derived from the detected acoustic signal;

determining a source of the detected acoustic signal to be unsure, if:

a prior determination of the source of the detected acoustic signal is that

the source of the detected acoustic signal was an audio transducer device;

the frame of audio data is classified as a live-type frame;

a predetermined number of most recent frames do not include enough live-type frames to exceed a predetermined live-type frame count threshold; and

an elapsed time since receiving a previous frame derived from speech exceeds a predetermined first time threshold; or

the prior determination of the source of the detected acoustic signal is that the source was the acoustic signal generated by a person speaking;

the frame of audio data is classified as a phone-type frame;

an elapsed time since receiving a previous live-type frame exceeds a predetermined second time threshold; and

a counter value is below a predetermined count threshold, the counter value to track a number of consecutive non-live-type frames received after receiving the live-type frame of most recent frames does not include enough live-type frames to exceed the predetermined count threshold;

determining the source of the detected acoustic signal to be the acoustic signal generated by the person speaking, if:

the prior determination of a source of a detected acoustic signal is unsure;

the frame of audio data is classified as the live-type frame; and

the predetermined number of most recent prior frames includes live-type frames that exceed in number the predetermined live-type frame count threshold;

determining the source of the detected acoustic signal to be the audio transducer device, if:

the prior determination of the source of the detected acoustic signal is unsure;

the frame of audio data is classified as the phone-type frame; and

the predetermined number of most recent prior frames includes phone-type frames that exceed in number a predetermined phone-type frame count threshold;

classifying the received frame of audio data based on spectral data of the received frame of audio data, the classifying comprising classifying the received frame of audio data as one of a plurality of predetermined frame types comprising the live-type frame; the phone-type frame; and an unsure-type frame, wherein live-type frames represent frames determined to be derived from acoustic signals generated by the person speaking, and phone-type frames represent frames determined to be derived from acoustic signals generated by the audio transducer device, wherein parameters used to classify frames include high band noise floor energy, low band noise floor energy, frame high band energy, frame low band energy, a ratio of the frame high band energy to the frame low band energy;

classifying the received frame of audio data from non-spectral data of the received frame based on a parameters energy ratio threshold for live speech and an energy ratio for phone speech; and

providing a signal indicating to a second computing device that the detected acoustic signal was generated by a person.

71. (Previously Presented) The computer-readable tangible medium of claim 70, wherein determining whether the detected acoustic signal was generated by a person further comprises:

determining whether the detected acoustic signal was speech from an audio transducer device.

72. (Previously Presented) The computer-readable tangible medium of claim 70, wherein the signal is transmitted over a network to participants of a multi-party conference.

73. (Previously Presented) The computer-readable tangible medium of claim 70, wherein determining whether the detected acoustic signal was generated by a person further comprises:

determining a source of a portion of the detected acoustic signal used to derive the frame based on the classification of the frame and a prior determination of a source of a detected acoustic signal portion.

74. (Previously Presented) The computer-readable tangible medium of claim 73, wherein the spectral data is obtained by performing a frequency transform on the frame of audio data.

75. (Previously Presented) The computer-readable tangible medium of claim 73, wherein the operations further comprise determining a first frequency band's energy and a second frequency band's energy from the spectral data.

76. (Previously Presented) The computer-readable tangible medium of claim 75, wherein the first frequency band corresponds to a frequency range for consonants and the second frequency band corresponds to a frequency range for vowels.

77. (Previously Presented) The computer-readable tangible medium of claim 76, wherein the operations further comprise:

classifying the frame as being generated by a person when the ratio of the energies of the first and second frequency bands exceeds a first predetermined threshold.

78. (Previously Presented) The computer-readable tangible medium of claim 77, wherein the operations further comprise:

selectively classifying the frame as being from another source when the ratio of the energies of the first and second frequency bands is below a second predetermined threshold.

79. (Previously Presented) The computer-readable tangible medium of claim 78, wherein the operations further comprise:

selectively classifying the frame as having an unknown source when the ratio of the energies of the first and second frequency bands exceeds the second predetermined threshold and is below the first predetermined threshold, the second predetermined threshold being less than the first predetermined threshold.

80. (Previously Presented) The computer-readable tangible medium of claim 76, wherein the operations further comprise:

determining whether the frame was derived from speech, wherein speech includes acoustic signals generated by a person speaking and acoustic signals generated by an audio transducer device.

81. (Previously Presented) The computer-readable tangible medium of claim 80, wherein the operations further comprise:

determining noise floor energies of the first and second frequency bands using the spectral data, wherein a frame is selectively classified as being derived from speech in response to the energy of the first frequency band exceeding the noise floor energy of the first frequency band or the energy of the second frequency band exceeding the noise floor of the second frequency band, or both.

82. (Previously Presented) The computer-readable tangible medium of claim 81, wherein classifying the received frame further comprises:

determining whether a frame adjacent to the received frame was classified as derived from speech.

83. (Previously Presented) The computer-readable tangible medium of claim 81, wherein classifying the received frame further comprises:

determining whether a frame within a predetermined number of frames relative to the received frame has substantially all of its energy in the second frequency band.

84. (Canceled)

85. (Previously Presented) The computer-readable tangible medium of claim 70, wherein the operations further comprise determining the source of the detected acoustic signal to be an acoustic signal generated by a person, if:

the prior determination of a source of a detected acoustic signal portion is that the source was an audio transducer device;

the frame is classified as a live-type frame; and

a predetermined number of prior frames includes live-type frames that exceed a predetermined live-type frame count threshold.

86. (Canceled)

87. (Previously Presented) The computer-readable tangible medium of claim 70, wherein the operations further comprise determining the source of the detected acoustic signal to be an audio transducer device, if:

the prior determination of a source of a detected acoustic signal portion is that the source was an acoustic signal generated by a person speaking;

the frame is classified as a phone-type frame;

an elapsed time since receiving a previous live-type frame exceeds a predetermined second time threshold; and

a counter value does not exceed a predetermined count threshold, the counter value to track a number of consecutive non-live-type frames received after receiving a

live-type frame of most recent frames do not include enough live-type frames to exceed a predetermined threshold.

88. (Canceled)

89. (Canceled)

90. (Canceled)

91. (Previously Presented) A system, comprising:

a first participant equipment including:

an audio discriminator to determine whether a detected acoustic signal was generated by a person in speaking; and

a sourcing signaler to selectively provide a signal indicating to a second participant equipment that the person is speaking in response to a determination by the audio discriminator that the detected acoustic signal was generated by the person in speaking;

wherein the audio discriminator comprises a spectrum analyzer to obtain spectrum data from a frame of audio data acquired from the detected acoustic signal and a frame classifier to determine which of a set of frame-types the frame matches, based on the spectrum data obtained by the spectrum analyzer;

wherein the frame types include a live-type frame, a phone-type frame, and an unsure-type frame, wherein the live-type frames represent frames

determined to be derived from acoustic signals generated by a person speaking, the phone-type frames represent frames determined to be derived from acoustic signals generated by an audio transducer device, and the unsure-type frames represent frames determined to be neither the live-type frames nor the phone-type frames.

92. (Original) The system of claim 91, wherein the sourcing signaler is to provide the signal to participants of a multi-party conference over a network.

93. (Original) The system of claim 91, wherein the audio discriminator is a component of the sourcing signaler.

94. (Previously Presented) The system of claim 91, wherein the audio discriminator comprises:

a hidden Markov model (HMM) machine to determine whether the detected acoustic signal was sourced by an acoustic signal generated by a person speaking using the classification of the frame-type.

95. (Original) The system of claim 91, wherein the audio discriminator comprises:

a spectrum analyzer to obtain spectrum data from a frame of audio data acquired from the detected acoustic signal;

a frame classifier to determine which of a set of frame-types the frame matches, based on the spectrum data obtained by the spectrum analyzer; and

a finite state machine (FSM) to determine whether the detected acoustic signal was sourced by an acoustic signal generated by a person speaking using the classification of the frame-type.